National Centers for Environmental Modeling
Environmental Modeling Center

DATA DOCUMENTATION
FOR
NOAA Operational Modeling Archive Distributed System
(NOMADS)

The Global Model and Cycling Analysis Rerun and Retrospective data set

DATA SET 6172

November 2001

Prepared for
the National Climatic Data Center (NCDC)

by

Jordan C. Alpert
NOAA/NWS
NCEP/EMC
5200 Auth Road
Camp Springs, MD 20746
(301) 763-8000x7205
http://www.ncep.noaa.gov

Table of Contents

TOPIC	PAGE
INTRODUCTION	
1. Data Set ID	3
2. Data Set Name	3
3. Data Set Aliases	3
DESCRIPTION	
4. Content of the Archived Data	3
5. Access Method for Archived Data	5
6. Element Names and Definitions	9
7. Start Date	9
8. Stop Date	9
9. Coverage	9
10. Location	9
11. Keyword	9
12. How to Order Data	10
DATA CENTER	
13. Archiving Data Center	10
14. Technical Contact	10
DATA QUALITY	
15. Known Problems	10
16. Quality Statement	10
17. Revision Date	10
OTHER DATA SETS	
18. Source Data Sets	10
19. Essential Companion Data Sets	10
20. Derived Data Sets	10
SUMMARIZATION	
21. References	11
22. Summary	11

1. **Data Set ID:**

6172

2. **Data Set Name:**

The Global Model and Cycling Analysis Rerun and Retrospective data archive

3. **Data Set Aliases:**

Global model analysis, Global Spectral Initial Conditions, Statistical Spectral Interpolation start-up files, Rerun or Retrospective files for analysis and model initialization.

4. **Contents of the Archived Data:**

The archived data set contains the minimum starting conditions for either the NCEP Operational Spectral Statistical Interpolation (SSI) cycling analysis and/or the Global Spectral Model (GSM). There are two types of files within the data set, Observations and restart files. The observation files are divided into BUFR files and IEEE files. The BUFR files are documented below. The remaining IEEE files are from remote sensing sources and are in the process of being converted to BUFR. We expect that only experts will access these files directly and this would be done through web based ftp services in the NOMADS system. However, a POST program for converting the model restart files from Spectral coefficients, on sigma vertical coordinate and gaussian grids to latitude/longitude, on standard pressure surfaces will be part of the NOMADS system.

The restart files are binary files that will be of interest to modelers or experts who want to obtain unchanged direct model results, that is, in the models own coordinate. Restart files will ordinarily transformed to pressure coordinate, longitude/latitude, GRIB grid data fields by a program called POST. The POST program is run and the server systems should present the GRIB grid data results transparently to users. A POST program will be supplied and documented elsewhere.

The documentation from the web page
http://www.emc.ncep.noaa.gov/mmb/papers/keyser/data_processing
by Dennis Keyser is shown below. It contains information on how to read BUFR files and details of the processing in NCEP Operations.

Observational Data Processing at NCEP
Dennis Keyser - NOAA/NWS/NCEP/EMC
(Last Revised 10/26/2001)

Most of the observational data at NCEP are stored in WMO BUFR format. This format is an international standard and provides an efficient means for transferring data. In addition it allows for great flexibility for adding new observation elements.

There are a series of tables associated with BUFR. Table A defines the data category associated with a particular BUFR message containing report data. Table B classifies and defines data

elements, or descriptors, according to scale, reference value, number of bits and units. Table C defines data description operators. Table D defines the list of common sequences. In addition, there are BUFR code and flag tables as well as code tables common to BUFR and other binary and alphanumeric codes. The need for external tables can make the process for BUFR data encoding and decoding quite cumbersome for a typical user.
As such, a special application has been designed at NCEP which provides user-friendly access to the BUFR files through a series of FORTRAN subroutines in a machine independent BUFR library (called BUFRLIB). These routines allow one to encode or decode data into BUFR using mnemonics to represent the data. The mnemonics are associated with BUFR descriptors in a special version of the Tables A, B and D (Table C is not yet included). When a BUFR file is created, the mnemonic table is read in from an external location and is itself encoded into BUFR messages at the top of the output file. These messages have Table A data category (message type) 11 (BUFR tables). This allows each BUFR file to be ?self defined?. No external tables are needed to decode data out of the file.

NCEP Central Operations has written a BUFRLIB software user guide which provides a detailed explanation of the NCEP BUFRLIB subroutines along with other useful information on BUFR as it is used at NCEP. (http://www.ncep.noaa.gov/NCO/DMQAB/Decoders/BUFRLIB/)

Next is a brief outline on the current method for processing observations that arrive at NCEP. Its main function is to provide links to web pages which discuss each item in detail.

1a. Raw bulletins arrive at NCEP on the GTS and are ingested into the database on the NCEP IBM-SP machine. Most of the data are stored in WMO BUFR format.
(http://www.ncep.noaa.gov/NCO/DMQAB/Decoders/)
- or --

1b. Most Satellite data are transferred directly from various NESDIS servers and then ingested into the BUFR database.
(http://www.emc.ncep.noaa.gov/mmb/papers/keyser/satellite_ingest.doc/document.htm)

2. Each NCEP network performs a time-windowed dump of data from the database at the scheduled data cutoff time.
(http://www.emc.ncep.noaa.gov/mmb/papers/keyser/data_dumping.doc/document.htm)

3. A series of programs process the observations into a monolithic, quality controlled file known as the "PREPBUFR" file.
(http://www.emc.ncep.noaa.gov/mmb/papers/keyser/prepbufr.doc/document.htm)

4. The PREPBUFR file is then read by the analysis codes.
(Global SSI analysis system documentation can be found at:
http://sgi62.wwb.noaa.gov:8080/RTPUB/research/jhtml/ssi3.html)

Here are some other links to my web sites devoted to data processing:
PREPBUFR Report Types used by Global/SSI (Aviation and Final) systems

PREPBUFR Report Types used by Global CDAS/reanalysis systems
PREPBUFR Report Types used by ETA/3DVAR (EDAS, Eta and NGM) systems
PREPBUFR Report Types used by upper-air Rapid Update Cycle (RUC-2 A and B) systems
Summary of Virtual Temperature Processing in PREPBUFR

RAOB/PIBAL Balloon Drift Processing in PREPBUFR

Sample program to decode reports from PREPBUFR file

Summary of Changes to the PREPDATA Program

BUFR Table B Descriptors and Mnemonics in NCEP Observational Database
Satellite Historical Documentation

NMC Office Note 29

NMC Office Note 124

5. **Access Method for Archived Data:**

The NCEP Global Spectral Model (GSM) files transmitted to NCDC represent the first operational model restart and retrospective archive. The entire data set of model run history is too large to keep at this time. Thus, we have devised the minimum set necessary to regenerate, as closely as possible, an operational run with the NCEP system and allow for other test beds to utilize, run experiments, and other models to initialize from this data set. The data set consists of conventional and remote sensed observations made ready to start NCEP cycling analysis ssytem. Other NCEP operational models in addition to the GSM, such as the Eta and WRF regional models are coming soon. Additional data sets consisting of model run history in GRIB pressure vertical coordinate on a longitude/latitude grid will be present as determined by NOMADS panel.

In addition, using the POST program, the restart files are converted to GRIB data sets. The information contained in restart (in this case GSM sigma files) files represent the final analysis of conventional and remote sensed observations in the models vertical sigma coordinate and spectral coefficients. This file is needed to rerun NCEP models and analysis. The POST program changes this file to standard WMO GRIB containing fields in pressure vertical coordinate on a longitude/latitude grid. The GRIB data set can be used as a verification set. The conventional and remote sensed observations including quality control are part of the minimum set to restart the analysis cycling system or to start models directly from theses initial conditions..

The analysis, initial condition and predicted fields on the model sigma levels are interpolated to the standard pressure levels in the POST program. The input file consists of the sigma level dependent variables and the output file consists of the pressure level variables in WMO standard GRIB. The file also contains several processed arrays (e.g. boundary layer parameters and tropopause parameters).

The Sigma file Contains (Table 1a) atmospheric variables on sigma surface and model sigma levels as well as topography. This file is the input to the POST program. Fixed fields are found in the Surface file (Table 1b). Normally, the Surface file and Sigma file are for internal use at NCEP and other designated test bed facilities for creating analysis/model reruns. The are made available to modelers and experts in this "raw" form if no interpolations or conversions are needed. We do not expect that these data sets will be directly accessed at the user level however, modelers and experimenters who wish to examine the analysis and initial conditions without any interpolations can obtain the data through web based ftp.

Sigma File (record number)	Contents	Length (bytes)	Type
	Table 1a: Structure of a Sigma File		
1	see NMC Office Note 85	32	binary
2	forecast hour	4	real
	initial hour	4	integer
	initial month	4	"
	initial day	4	"
	initial year	4	"
	sigma levels (1)	(KDIM + 1) x 4	real
	sigma layers (2)	KDIM x 4	"
3	Orography in meters (spherical coefficients)	MDIM x 4	"
4	Spherical coefficients of ln (p_s) , where p_s is surface pressure (cb)	"	"
5-22	Temperature (°K) in model layers 1-KDIM (spherical coefficients)	"	"
23,24	Divergence and Vorticity alternating through layer 1....	"	"
...57,58	....thru layer KDIM	"	"
59-70	Specific humidity in model layers 1-KDIM (spectral coefficients)	"	"

Note all the spherical coefficients are stored in this order: real part, imaginary part, N-S index and E-W wavenumber.

1) Sigma levels are the level starting from $\sigma = 1$ at the surface and ending at $\sigma = 0$ at the top. Only derived quantities, (vertical velocity, various fluxes) are defined on these levels.

2) Sigma layers are where dependent variables (T, D, ζ , u, v, q) are defined

Table 1b: Structure of a Surface File

Surface File (record number)	Contents	Length (bytes)	Type
1	see NMC Office Note 85	32	binary
2	Forecast hour	4	real
	Initial hour	4	integer
	Initial month	4	"
	Initial day	4	"
	Initial year	4	"
3	Surface temperature	IDIM x JDIM x 4	real
4	Soil wetness	"	"
5	Snow depth	"	"
6	Sub-surface temperature , layer 1 (TG1)	"	"
7	Sub-surface temperature , layer 2 (TG2)	"	"
8	Sub-surface temperature , layer 3 (TG3)	"	"
9	Surface roughness length	"	"
10	Surface background albedo(1)	"	"
11	Surface-type mask(2)	"	"
12	High cloud fraction	"	"
13	Middle cloud fraction	"	"
14	Low cloud fraction	"	"

Note: All are gaussian gridded arrays of IDIM x JDIM, where I=1 is 0°E (then eastward) and J=1 is near the North Pole (then southward).

(l) Albedo is the background albedo that is modified by snow cover.

(2) Ocean = 0., land = l., and sea ice = 2.

6. **Element Names and Definitions:**
See Office Note 388 and its supplement (1998) Table 2 “Parameters and Units” pp 45-52.

7. **Start Date:** To be determined

DATA AVAILABILITY: To be determined

10. **Location:**

a. Global

11. **Keywords:**

12. **How to Order Data:**

The cost for this data when accessed through NOMADS system servers or associated ftp web based services is free. For more information contact:

National Climatic Data Center
151 Patton Avenue
Asheville, NC 28801-5001

Phone 828-271-4800
FAX 828-271-4876
e-mail orders@ncdc.noaa.gov

13. **Archiving Data Center:**

National Climatic Data Center
151 Patton Avenue
Asheville, NC 28801-5001

14. **Technical Contact:**

Jordan C. Alpert
NOAA/NWS/NCEP/EMC
5200 Auth Road, Rm. 207
Camp Springs, MD 20746
jordan.alpert@noaa.gov
(301) 763-8000 x7205
http://www.ncep.noaa.gov

15. **Known Uncorrected Problems:**

16. **Quality Statement:**

17. **Revision Date:**

23 November 2001

18. **Source Data Sets:**

19. **Essential Companion Data Sets:**

20. **Derived Data Sets:**

21. **References:**

Kanamitsu, M., 1989: Description of the NMC Global Data Assimilation and Forecast System, Weather and Forecasting, 4(335-342).

Sela, J.G., 1980: Spectral modeling at the National Meteorological Center, Mon. Wea. Rev., 108 (1279-1292).

22. **Summary:**

The National Weather Service's National Center for Environmental Prediction (NCEP) runs a series of computer analyses and forecasts operationally. One of the primary operational systems is the Global Data Assimilation System (Kanamitsu, 1989), which uses the spectral Medium Range Forecast model (MRF) for the forecast (Sela, 1980). In simple terms, for each run, unequally spaced conventional and remote sensed observations are assimilated with "first guess" data fields (forecasts from the previous model run), and dynamic imbalances in the data are reduced, resulting in "analyzed" data fields. Then the forecast is made. The analyzed data provides an optimal representation of the real atmosphere on a grid or spectral coefficients of spherical harmonics. These can be compared to observations which have limitations due to measurement error or other instrument problems, and nonuniform spatial and temporal distributions of the observations.

ORIGIN OF DATA

The enclosed rerun/retrospective source data sets contains the NCEP operational ready observation and restart files necessary to

begin the Spectral Statistical Interpolation (SSI) cycling analysis. The collection of observations, conventional and remote sensed with the previous model forecasts which are used as a "guess" and restart files can be used to rerun the analysis cycling or global forecast system.

NCEP post-processing of model run history using the POST program is used to convert the restart files to pressure coordinate, longitude/latitude GRIB grids.

The archiving program extracts 4 times a day minimum restart file set from the NCEP operational communication directory. In particular, each data set contains:

1) Conventional and remote sensed observations acted upon by the NCEP anlaysis SSI program which exercise quality control, windowing, filtering, etc... at 00Z, 06Z, 12Z and 18Z,

2) The model restart files (sometimes called siganl and sfcanl) are present in the archive only at 00Z.

From 2) one can start the model and integrate for 6 hours creating a 6 hour "guess" which with 1) can be used to start the SSI to create the next analysis file at that time. The 6 hr analysis file in turn can (re-)start the model creating the next 6 hr "guess" and so on creating initial conditions. In this way the data set represents the minimum set of Observations and restart files to reconstruct, as closely as possible, the NCEP Operational (in this case) global model suite. The data set names and size (bytes) are in Table 2. At a time when disk space is a premium comparted to CPU resources, the minimum set represents a way to provide model data sets in their original forms as well as post processed fields and their components.

Table 2: List of files and file size of the minimum data set for the NCEP SSI and GSM. There are links to a few files because of past name usage. The size of the data set is 370Mb.

YYYY is the year
MM is the month
DD is the day

8279556	a1bn15.YYYYMMDD00
4285516	a1bn15.YYYYMMDD06
8275464	a1bn15.YYYYMMDD12
8245952	a1bn15.YYYYMMDD18
4865340	a1bn16.YYYYMMDD00
6552112	a1bn16.YYYYMMDD06
8427612	a1bn16.YYYYMMDD12

```
 5999196  a1bn16.YYYYMMDD18
 8349304  adpsfc.YYYYMMDD00
 8290720  adpsfc.YYYYMMDD06
 8271520  adpsfc.YYYYMMDD12
 8091792  adpsfc.YYYYMMDD18
 1184696  adpupa.YYYYMMDD00
  161904  adpupa.YYYYMMDD06
 1166256  adpupa.YYYYMMDD12
  136472  adpupa.YYYYMMDD18
 2337856  aircar.YYYYMMDD00
 1625448  aircar.YYYYMMDD06
 1771112  aircar.YYYYMMDD12
 2129384  aircar.YYYYMMDD18
  651560  aircft.YYYYMMDD00
  747808  aircft.YYYYMMDD06
 1025112  aircft.YYYYMMDD12
  936664  aircft.YYYYMMDD18
10083900  b1bn15.YYYYMMDD00
 5174856  b1bn15.YYYYMMDD06
10242828  b1bn15.YYYYMMDD12
10057440  b1bn15.YYYYMMDD18
 6010992  b1bn16.YYYYMMDD00
 8086212  b1bn16.YYYYMMDD06
10319436  b1bn16.YYYYMMDD12
 7361712  b1bn16.YYYYMMDD18
    7668  biascr.YYYYMMDD00
      19 gdas1.t00z.bufprepda -> gdas1.t00z.prepbufr
19256208 gdas1.t00z.prepbufr
      17 gdas1.t00z.sanl -> siganl.YYYYMMDD00
      17 gdas1.t00z.sfcanl -> sfcanl.YYYYMMDD00
 4830968  goesnd.YYYYMMDD00
 8073424  goesnd.YYYYMMDD06
 8217080  goesnd.YYYYMMDD12
 9033792  goesnd.YYYYMMDD18
 5827488  h1bn14.YYYYMMDD00
 4626384  h1bn14.YYYYMMDD06
 5693424  h1bn14.YYYYMMDD12
 5620848  h1bn14.YYYYMMDD18
 5492400  h1bn15.YYYYMMDD00
 2857776  h1bn15.YYYYMMDD06
 5556048  h1bn15.YYYYMMDD12
 5577648  h1bn15.YYYYMMDD18
 3240240  h1bn16.YYYYMMDD00
 4373376  h1bn16.YYYYMMDD06
 5637840  h1bn16.YYYYMMDD12
 3952032  h1bn16.YYYYMMDD18
```

```
  226884  icegrb.YYYYMMDD00
  226884  icegrb.YYYYMMDD06
  226884  icegrb.YYYYMMDD12
  226884  icegrb.YYYYMMDD18
  770032  m1bn14.YYYYMMDD00
  606272  m1bn14.YYYYMMDD06
  772352  m1bn14.YYYYMMDD12
  733632  m1bn14.YYYYMMDD18
  309800  osbuvb.YYYYMMDD00
  284824  osbuvb.YYYYMMDD06
  344760  osbuvb.YYYYMMDD12
  327720  osbuvb.YYYYMMDD18
  160120  proflr.YYYYMMDD00
  157040  proflr.YYYYMMDD06
  159648  proflr.YYYYMMDD12
  161168  proflr.YYYYMMDD18
  276192  qkswnd.YYYYMMDD00
  270816  qkswnd.YYYYMMDD06
  192416  qkswnd.YYYYMMDD12
  350888  qkswnd.YYYYMMDD18
 7662616  satwnd.YYYYMMDD00
 5937184  satwnd.YYYYMMDD06
 6222816  satwnd.YYYYMMDD12
 9127080  satwnd.YYYYMMDD18
  210060  sbvn16.YYYYMMDD00
  252072  sbvn16.YYYYMMDD06
  294084  sbvn16.YYYYMMDD12
  252072  sbvn16.YYYYMMDD18
13632240  sfcanl.YYYYMMDD00
   35040  sfcbog.YYYYMMDD00
       0  sfcbog.YYYYMMDD06
   35872  sfcbog.YYYYMMDD12
       0  sfcbog.YYYYMMDD18
 1417632  sfcshp.YYYYMMDD00
 1341008  sfcshp.YYYYMMDD06
 1494528  sfcshp.YYYYMMDD12
 1541848  sfcshp.YYYYMMDD18
29885672  siganl.YYYYMMDD00
  714948  snogrb.YYYYMMDD00
  714948  snogrb.YYYYMMDD06
  714948  snogrb.YYYYMMDD12
  714948  snogrb.YYYYMMDD18
 2459168  spssmi.YYYYMMDD00
 2431384  spssmi.YYYYMMDD06
 2518328  spssmi.YYYYMMDD12
 2514872  spssmi.YYYYMMDD18
```

326408 sptrmm.YYYYMMDD00
331688 sptrmm.YYYYMMDD06
332656 sptrmm.YYYYMMDD12
324368 sptrmm.YYYYMMDD18
154068 sstgrb.YYYYMMDD00
154068 sstgrb.YYYYMMDD06
154068 sstgrb.YYYYMMDD12
154068 sstgrb.YYYYMMDD18
12933 stat01.YYYYMMDD00
13104 stat01.YYYYMMDD06
12933 stat01.YYYYMMDD12
13105 stat01.YYYYMMDD18
3142 stat02.YYYYMMDD00
3142 stat02.YYYYMMDD06
3142 stat02.YYYYMMDD12
3142 stat02.YYYYMMDD18
7449 statup.YYYYMMDD00
7854 statup.YYYYMMDD06
7449 statup.YYYYMMDD12
6963 statup.YYYYMMDD18
0 tcvitl.YYYYMMDD00
0 tcvitl.YYYYMMDD06
0 tcvitl.YYYYMMDD12
0 tcvitl.YYYYMMDD18
345360 vadwnd.YYYYMMDD00
342248 vadwnd.YYYYMMDD06
325312 vadwnd.YYYYMMDD12
346232 vadwnd.YYYYMMDD18